Липницкая, О. Л Модернизация исследовательского труда историка с помощью информационных технологий / О. Л. Липницкая // Працы гістарычнага факультэта БДУ: Навук. зб. Вып. 4 / Рэдкал.: У. К. Коршук (адк. рэд.) [і інш.] — Мн: БГУ, 2009. — С. 161–165.

О. Л. ЛИПНИЦКАЯ

МОДЕРНИЗАЦИЯ ИССЛЕДОВАТЕЛЬСКОГО ТРУДА ИСТОРИКА С ПОМОЩЬЮ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Использование достижений современных информационных технологий позволяет использовать нетрадиционные источники информации, развивает эффективность исследовательской работы. На возможности, которые представляют информационные технологии историкам, указывал еще И. Д. Ковальченко: «Математические методы и ЭВМ позволяют существенно углубить исторические исследования, сделать их более объективными и доказательными. Ведь раскрытие количественной меры качественно содержательной определенности любых явлений естественного и общественного мира является высшим уровнем научных исследований» [1, с. 4].

В теории источниковедения исторические источники в зависимости от критерия ранжированы на виды, типы, роды, т. е. существуют разные классификации. И соответственно исследователь выбирает те или иные методы и приемы для изучения источника.

Есть классификация исторических источников, которая исходит из того, в каком формате представлен исторический электронный источник и по какому принципу смоделирована информация в нем [2, с. 319—326]. В соответствии с этими критериями исторические источники разделяются на статистические, структурированные, текстовые или нарративные, изобразительные и картографические. Для каждой из этих групп определены свои методы и приемы исследования — компьютерные программные средства, которые и могут стать инструментарием исследования.

Конечно, многообразие исторических источников, неисчерпаемая информация, заключенная в них, являются причиной того, что ни одна из существующих программ не может обеспечить исследователя всеми необходимыми средствами анализа. Каждая из программ воспринимает информацию определенного вида (типов) и интерпретирует ее своим строго определенным образом. На сегодняшний день существует определенный минимум программного обеспечения, который позволяет историкам, архивистам быть «вооруженными» знанием компьютерных технологий как для изучения исторических источников, так и для исторического исследования.

Так, одним из оптимальных программных средств для работы со статистическими источниками являются электронные таблицы (ЭТ).

Тот факт, что информация в статистическом источнике имеет табличное представление, позволяет сохранить его практически в неизменном виде при переводе в электронный вид посредством, например, MS Excel. Аналитические возможности ЭТ полностью отвечают потребностям историков, использующих традиционные методы, в первичной обработке статистического материала. Привлекательны для исследователя и возможности графического представления как исходных данных, так и результатов их обработки. Кроме того, возможен и более сложный анализ источников, в том числе и с помощью статистических пакетов, так как электронные таблицы совместимы с ними. Достаточно часто историки сталкиваются с проблемами точности и сопоставимости данных статистического источника и вынуждены решать их. Электронные таблицы не только легче решают такие проблемы, но и могут частично или полностью автоматизировать их решение.

Так, например, при изучении инвентаря имения можно составить по его данным таблицу и попытаться выяснить, какова была зависимость между размером крестьянского надела и количеством взрослых и несовершеннолетних сыновей у хозяина посредством вычисления коэффициента ковариации [3]. Посредством описательной статистики (одна из многочисленных функций, предоставляемых ЭТ) можно проанализировать данные таблицы «Партизаны Беларуси» [4], а затем представить полученные вторичные данные на графиках. Гистограммы, построенные ЭТ по годам, дают пики увеличения количества ставших партизанами в определенные месяцы — интерпретировать их появление несложно.

Структурированные источники можно адекватно анализировать с помощью технологий баз данных (БД). Так, еще в 1990-е гг. в Институте истории НАНБ на этапе подготовки к переизданию Пописов войска ВКЛ (33 книги РИБ, издание 1915 г.) были созданы базы данных [5], на основе которых решены такие практические археографические вопросы, как составление алфавитного именного и географического указателей.

Для анализа информации, содержащейся в тарифе подымного налогообложения Оршанского повета 1667 г., использованы новейшие компьютерные технологии. В качестве основной была избрана источнико-ориентированная стратегия, которая предполагает максимально бережное сохранение структуры и текста первичного источника. Основными этапами этой стратегии являются: 1) транслитерация источ-

ника; 2) перевод источника в машиночитаемый вид, создание компьютерной базы данных; 3) анализ количественных данных и их интерпретация.

Система управления базами данных является мощным аппаратом для решения многих исследовательских задач, поскольку позволяет выполнять самые разнообразные запросы исследователя, осуществлять поиск и выборку нужных данных. Полученные путем запроса данные могут быть представлены в виде диаграмм и таблиц, а также подвергнуты последующему статистическому анализу.

В результате проведенного в Институте истории НАНБ анализа данных тарифа Оршанского повета [6] получены ответы на следующие вопросы: каковы были последствия войны 1548—1667 гг.; каков был социальный состав налогоплательщиков; каков был типичный налогоплательщик Оршанского повета 1667 г.; выявить взаимосвязь сословия и размера недвижимой собственности; характер прав на недвижимую собственность; степень имущественной дифференциации налогоплательщиков; определить виды и типы недвижимой собственности; сгруппировать налогоплательщиков по величине земельных владений.

Использование в своих исследованиях только реляционной модели базы данных заставляет историков, работающих со сложными текстовыми источниками, пожертвовать целостностью или первоначальной оригинальностью текста. Ведь в значительной части текстовых исторических источников данные настолько «рассыпаны» и сложны, что требуются специфические инструментальные средства, а именно специальное программное обеспечение. В силу этой специфичности источников для их анализа используются новые пакеты, как например разработанная в Германии система KLEIO или система TACT — продукт Торонтского университета.

Текст источника, коллекции источников переводятся в электронный вид посредством сканирования или непосредственного ввода через клавиатуру. Текстовый файл преобразуется в текстовую базу данных, структурируется и маркируется с помощью описательных языков разметки документа, предоставляемых полнотекстовой системой TACT.

Программа была разработана для лингвистического анализа текста, схема работы предполагает осуществление трех последовательных процедур. В ходе первого этапа в тексте документа выявляется ряд терминов, характеризующих тематику исследуемого документа. На их основе формируются обобщенные понятия (смысловые категории), наиболее точно и полно описывающие основные смысловые блоки текста. Итог

первого этапа — составление (одной или нескольких) исследовательских баз данных. На втором этапе происходит подсчет частот встречаемости категорий в тексте, анализ связей, выявление контекстов употребления в документе. Строятся таблицы и графики смысловых взаимосвязей. На заключительном этапе полученные данные интерпретируются исследователем. То есть происходит контент-анализ источников.

Исторические процессы, факты, события происходят не только во времени, но и в пространстве. Новейшие компьютерные технологии дают возможность не только переводить в электронный вид картографические источники прошлого, но и создавать компьютерные карты по данным письменных источников, включающих такого рода информацию.

Геоинформационная карта, полученная с помощью ГИС (географическая информационная система), обладает возможностями выводить на карту часть информации согласно запросам, комбинировать различные данные, соединяться с базами данных и пр. Иными словами, компьютерная карта — это средство моделирования различных процессов в пространстве и времени. Компьютерная карта, по существу, представляет собой сочетание двух видов данных — географической и статистической информации.

Так, для создания компьютерной карты населенных пунктов на основе архивного источника, например «Тарифа … Менского повета 1667 г.», в качестве карты — «растровой подложки» можно использовать карту «Великое княжество Литовское в конце XVI в.» М. Ф. Спиридонова и/или современную топографическую карту Беларуси в масштабе 1:200 000. Историческая информация, которую можно представить на карте, кроме локализации самих населенных пунктов, может быть принадлежность к владельцам, ранжировка в зависимости от количества дымов, очерчивание границ имений, группировка по типам владений. Первым шагом в составлении электронной карты является их локализация. На завершающем этапе после интеграции информационных слоев в единую карту и ее оформления можно создавать тематические карты, производить анализ данных, делать запросы и исторические выводы.

Таким образом, используя известные и новые компьютерные технологии и применяя их к разным историческим источникам, как к рукописным, так и опубликованным, можно успешно переводить их в электронный вид, обрабатывать и анализировать, ставить и решать исторические задачи.

ЛИТЕРАТУРА

1. Цит. по: К 10-летию Информационного бюллетеня АИК // ИБ АИК. № 25. М., 2000.

2. *Белова, Е.* Историческая информатика / Е. Белова, Л. И. Бородкин, И. М. Гарскова. М., 1996.

3. Фрагмент инвентаря имения Везовец Слонимского повета Новогородского воеводства, датированный 22 апреля 1583 г., опубликован М. Ф. Спиридоновым в «Навінах АН БССР».

4. Автор выражает благодарность Е. Я. Павловой за разрешение использовать фрагмент таблицы, составленной на основе архивных дел из фондов Национального архива РБ; № 3500, № 3706—3712, № 4052, № 4053, № 4058—4080.

5. БД создана на основе научно-исследовательской базы «Пописы войска Великого княжества Литовского 1528 и 1567 гг.», разработанной в Институте истории НАНБ при участии автора.

6. БД создана на основе научно-исследовательской базы «Тариф подымного налогообложения Оршанского повета 1667 года», разработанной в Институте истории НАНБ при участии автора.